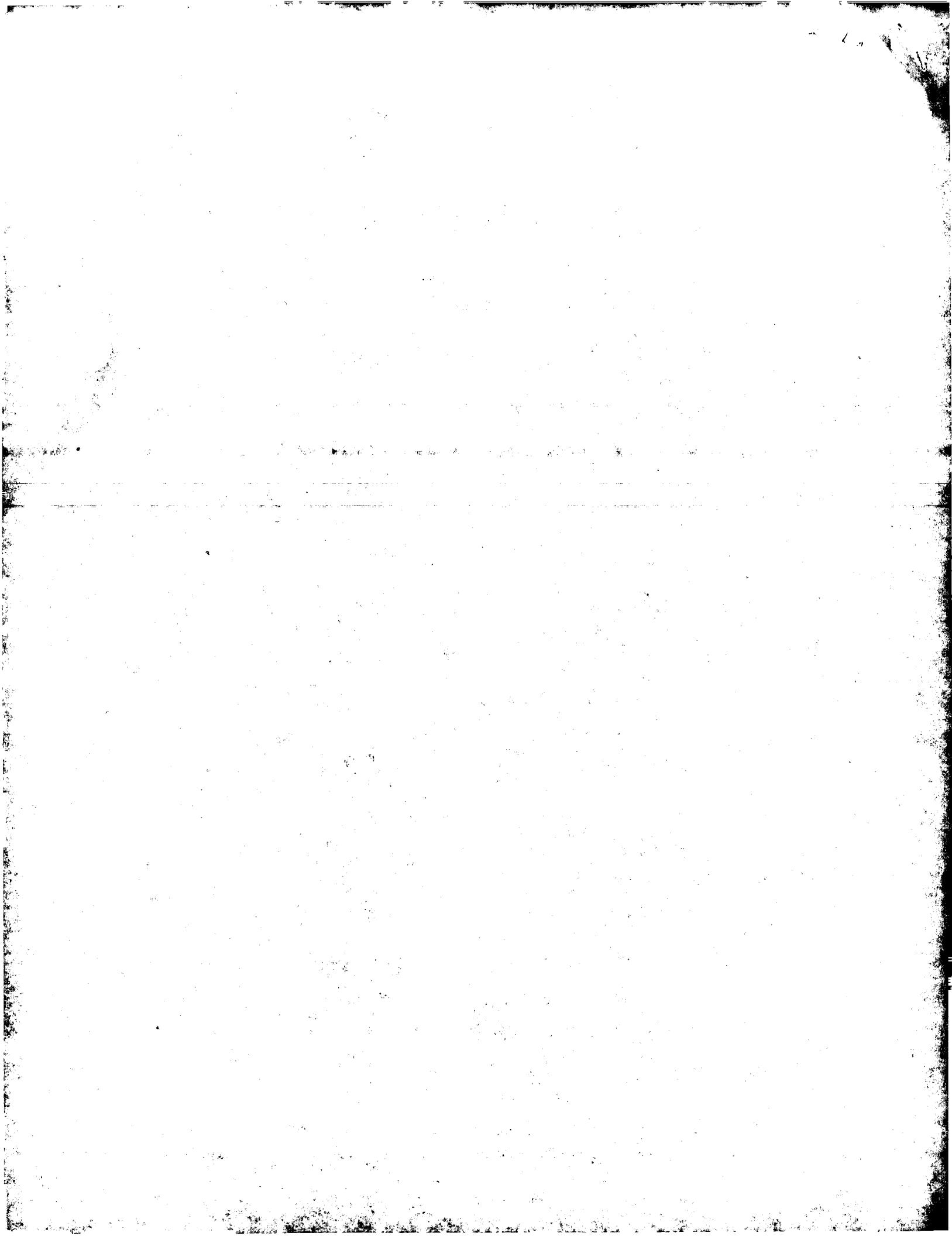


1 DNA_SEQUENCE 1.0
 LOCUS HSM00235 1433 bp mRNA linear PRI 18-FEB-2000
 DEFINITION Homo sapiens mRNA; cDNA DKF2p586B1922 (from clone DKF2p586B1922).
 ACCESSION AL049450
 VERSION AL049450.1
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
 Mammalia; Eutheria; Primates; Catarhini; Hominidae; Homo.
 REFERENCE
 AUTHORS Koehler, K., Beyer, A., Mewes, H. W., Gassnerhuber, J. and Wiemann, S.
 TITLE Direct Submission
 JOURNAL Submitted (10-MAR-1999) MIPS, Am Klopferspitz 18a, D-82152
 COMMENT
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email: S.wiemann@dkfz-heidelberg.de; sequenced by BMFZ (Biomedical Research Center at the Charite', Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.
 This clone (DKF2p586B1922) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://www.mips.biochem.mpg.de/proj/cDNA/>.
 FEATURES
 source
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKF2p586B1922"
 /tissue_type="uterus"
 /clone_id="586 (synonym: hutel)". Vector psp01; host DH10B; sires NotI + SalI/MluI"
 /dev_stage="adult"
 1386..1391.
 polyA_signal
 /polyA_site 1412 321 c 331 g 367 t
 BASE COUNT
 ORIGIN
 AL049450 Length: 1433 July 21, 2003 13:18 Type: N Check: 450 ..
 1 TAGGAAGAAG CGCTCAACAC AGAACACCC TAACTATCCTT ACTCCCTTC
 51 AGGACTAA TCCCTTCAGA GATTCCTT GCTTCAGTCCTC TTTCAGGAGCC
 101 TCCGGAAGAT GCTAGGAGTC GTTAAAGAG GCTGGAAGAT GCCCGTGG
 151 AGGAAACTGT TGGTAGCTGG CAATACTTCA TCCAAAGGAG GGGTGC
 201 TGCCTCCCTG GGCAGCAGAC ATCATGGATT TTCTCCCTC TGGAAACAG
 251 AGGTCCCTC GTTTCGGTT TGAATCCCTT AGAGACAGA AACTCTTC
 301 TCTCCCTAC CGCGCTGGGA CTGCTGGGA GGAGAGGGGG TGGAGGGTT
 351 AGACTCTCTAC CCCCATTTGG AAGTACAGAA AGAAAGACTC CTGTTAAAAA
 401 GCTGGCCCTT GAATTCTAA AACCTCAAAAC AAAAACCTT AAGCCTCTC
 451 CTATTAAGA GGAACCTGT GGTTCCTTAT CTGAAACCTGT TTGTAACGT
 501 GATGAGGA GCCAGAAC CCCAGAAAG CCCGGCTT CAGTGGACAC
 551 CCCACCAAGA CTCTCACTC CCCAAAGGG ACCCAGCAC CATTCCAGG
 601 AGAAAGCCTT CTCAGTGAG ATAGAAGATT TGGCGTACCT TTCCACCA
 651 GAAATGATT TGTGTGTG GCACCAAGCTT CCCCATCAC CGTTACATT
 701 AGGGATCC TCTCCAAGA AGGAGGAGC TGTAGCAAGT AAGGCA
 751 GAAACATTGC TCTTATACCC TAGTGGGGC GGTCAAGCTA ACAAGCTGA
 801 AAATGCCTT GGCATTTTA AAAAGTGCA ATCAATTAAG CAGAGCTCTG
 851 TCAAGATGA GTAAGTTAAC AGCCAGAGAC AGACACTGTG CAGGATG
 901 AATAGATGG AATTACAGCA AAATGTGCTC AATGTATTG CCTGGTTACA
 951 ACATGGAG ATGTGTTGCA CAGTAAGTGT CTCATCACAA GAGCACCAGA
 1001 CTGGGGGTG TAATCTCCGG CAACTGAT GCCCTCTGAA AGAAGGGTT
 1051 TCTGNGCTGT GAAATGCATA GAACTATACT TTGCCATGCA CGACGUTTC
 1101 TGCAATTGAT ATGTTGTAATGAGG GGGTCTTGT GGTGTTCTCA
 1151 GGGCCAAATG GAAATTG GGTGGGGAG CCAGCTGG GTGGGGATT
 1201 TTACCTCTGG CCGCCGCTCT TTAACTAT AACATTTAT CTGTAATCT
 1251 ATGTCCTGT CTGGGGGCA GGAGGAATCT GCCAAAGACC AACAGCTTA
 1301 CTTATCTA CTATACTCA CAAGGATCT AAAATGAA GAGTGTACTT
 1351 GGATTCGAGT AGCCCATGG TTGTCTCAT ATTAAATAA AATGGCTAC
 1401 AACTATTT TCAACAAAA AAAAAAAA AAA





1 mRNA-SEQUENCE 1.0
 LOCUS HSM0235 1433 bp mRNA linear PRI 18-FEB-2000
 DEFINITION Homo sapiens mRNA; cDNA DKFZp586B1922 (from clone DKFZp586B1922).
 ACCESSION AL049450
 VERSION AL049450-1
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 1433)
 AUTHORS Koehrer,K., Beyer,A., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
 COMMENT Direct Submission
 Submitted (10-MAR-1999) MIPS, Am Klopferspitz 18a, D-82152
 Martinsried, GERMANY
 Please contact the RZPD, Ressourcenzentrum, Heinberweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
 information about the clone and the sequencing project is available
 at <http://www.mips.biochem.mpg.de/proj/CDNA/>.
 FEATURES Location/Qualifiers
 source 1..1433
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="RZPD:DKFZp586B1922"
 /db_xref="Taxon:9606"
 /clone="DKFZp586B1922"
 /tissue_type="uterus"
 /clone_lid="386 (synonym: hutel). Vector pSport1; host
 DH10B; sites NotI + SalI/MluI"
 /dev_stage="adult"
 /polyA_signal 1386..1391
 polyA_site 1412
 BASE COUNT 414 a 321 c 331 g 367 t
 ORIGIN
 AL049450 Length: 1433 July 21, 2003 13:18 Type: N Check: 450 ..
 1 TAGGAGAAG CGCTCACAC AGAACACCT AAGTATCCCT ACTCCCTTC
 51 AAGGACTAA TCCCTTCAGA GATTCCTT GCTTCAGTC TTCAGGAGC
 101 TCCGAGAT GCTAGGAGTC GTTTAAGAG GCTGGAGAT GGCCGTTGG
 151 AAGGAAGTGT TGGTAGCTGG CAAACTCTCA TCCCAAGGAG GCGGTGGCAA
 201 TAGCTCCCTG GGCAGCAGAC ATCATGGATT TTCCCTCCCTC TGGGAACCAAG
 251 AGGTGCTCTC GTTTCTGGTT TGATCCCTT AGAGACAAG AACTTCATC
 301 TCTCCGATAC CGCGTGGGA CTGCTGGGA GGAGAGGGG TGGAGGGTT
 351 AGACATCTAC CCCATTGG AACATCAAAAC AAAAACCTCTT AAGCACTCTC
 401 GCTGGCTCTC GAATTTCAAAGG ACCCAGGAC CATCCAAAGG
 451 CTATTAAGA GGAAACCTGT GGTTCTTAT CTGAAACTGT TTGTAACAGT
 501 GATTGAGGA GCCAGAAC CCCAGAAAG CCCGGCTT CAGTGGACAC
 551 CCCACCAAGA CTCRCCACTC CCCAAAAGGG ACCCAGGAC CATCCAAAGG
 601 AGAAAGCCTT CTCAAGTGTG ATAGAGAGTT TGGCTACCT TTCCACCA
 651 GAAATGATT TTGTGCGTGC GCACCAAGCTT CCCCATCAC CGTACCAATT
 701 AGGGATATCC TCTCCAAGAA AGGAGGAGAC TGTAGCAAGT AAGGCAATAGA

751 GAACACTTGCTCTTATACCC TAGTGGTGC GCTCAAGCTA ACAAGSTGTA
 801 AAATGCTT GGATTTTA AAAAGCTCA ATCAATTAAG CAGAGTCTG
 851 TCAGAAATGA GTAAAGTTAAC AGCCAGGAC AGACAGTGT CAGGATTGC
 901 AAATGATGG AATTAACAGCA AAATGTCCTC AATGTAATGT CCTGCTTACA
 951 ACATGGGAG ATGTTGTCG CAGTAAGTGT CTCACTACAA GAGCACCAGA
 1001 CTTGGGGTG TAATCTCGG CAACTTCAT GCCCTCTGAA AGAAAGGGTT
 1051 TCTGTCGTGT GAATGCTA GAACTATACT TGGCCATGCA CGACGTCTC
 1101 TGCATTGAT ATGTTGTAATCTGGAG GTGGTCTTG GGTGTCCTCA
 1151 GGGSCCAATG GTAAATTG GGTTGGGAG CCAGCTGG GTGGGGAATT
 1201 TTCACTGAG CCTCGCTCTT TAACTATAT AACATTTAT CTGTAATCT
 1251 ATGTCCTGT CTGGGGSCA GGAGGAATCT GCAAGAGACC AACAGCTTA
 1301 CTTATCTA CTATACTCA CAAGAGTCTT AAAATGTGAAG GAGTTACTT
 1351 GGATGCACT AGCCATGG TTGTTCAT ATTAAATAA ATGGCTAC
 1401 AACATTTT TCAACACAAA AAAAAGAAA AAA



GenCore version 5.1.6
 Copyright (c) 1993 - 2003 CompuGen Ltd.

Run on: July 21, 2003, 13:29:30 ; search time 3 seconds
 (without alignments)
 4.177 Million cell updates/sec

Title: OM nucleic - nucleic search, using sw model
 Perfect score: HSM800235
 Sequence: 1 TGGAAAGAGCGCTAACAC... ACAAACAAACAAACTCTAGCAGCTC
 Scoring table: IDENTITY.NUC
 Gapop 10.0 , Gapext 0.5
 Searched: 1 seqs, 4372 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
 Maximum DB seq length: 200000000
 Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 1 summaries

Database : US10054935.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

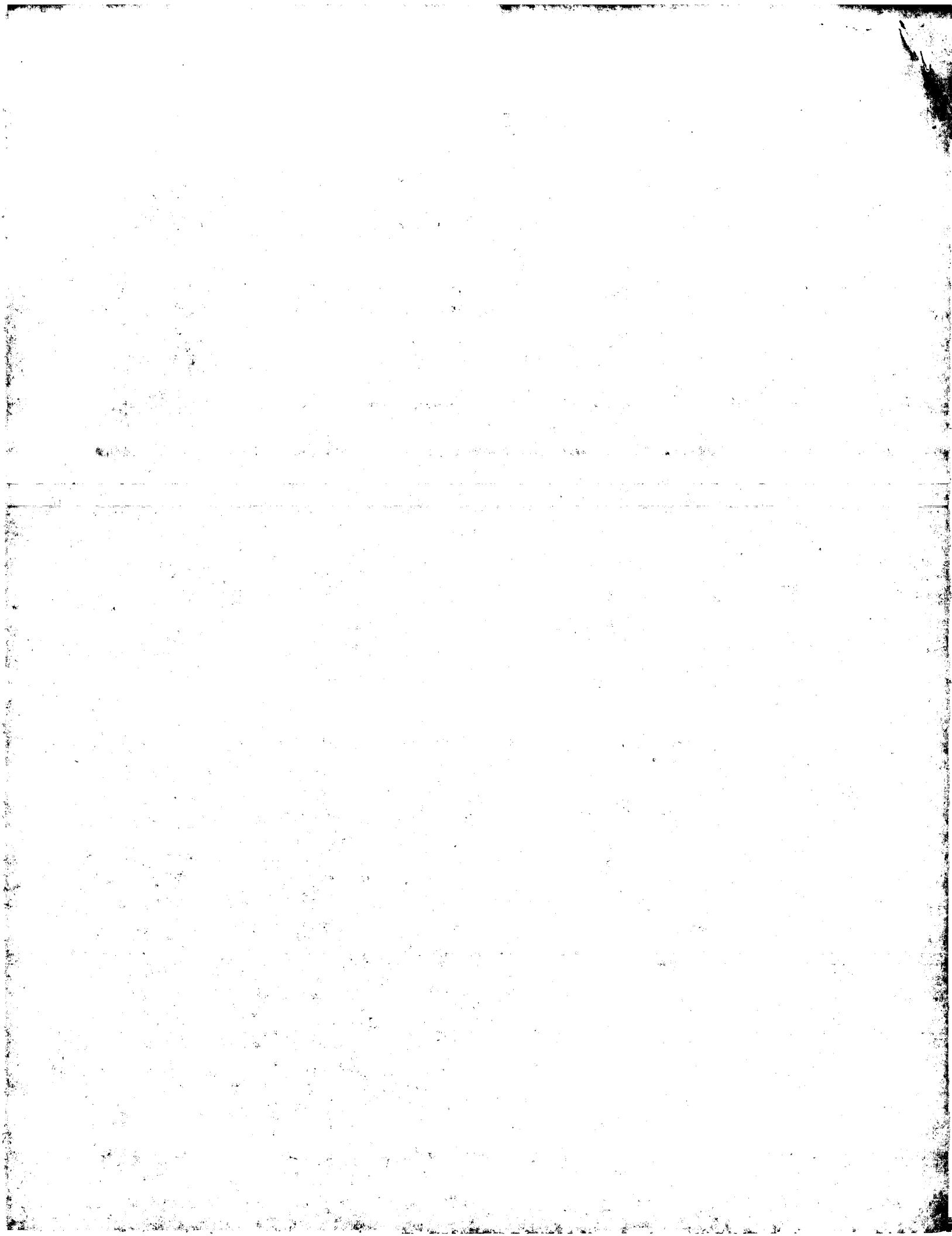
Result No.	Score	Query Match	Length	DB ID	Description
1	380.2	26.5	4372	1	US-10-054-935-1 Sequence 1, Appli

ALIGMENTS

RESULT 1
 US-10-054-935-1
 ; Sequence 1, Application US/10054935
 ; GENERAL INFORMATION:
 ; APPLICANT: Origene Technologies, Inc
 ; TITLE OF INVENTION: BREAST CANCER TRANSCRIPTION FACTOR GENE AND USES
 ; FILE REFERENCE: 160 107 R1
 ; CURRENT APPLICATION NUMBER: US/10/054,935
 ; CURRENT FILING DATE: 2002-01-25
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 4372
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (78)..(1922)
 ; OTHER INFORMATION:
 ; US-10-054-935-1

Query Match 26.5% ; Score 380.2; DB 1; Length 4372;
 Best Local Similarity 95.6%; Pred. No. 0; Mismatches 18; Indels 0; Gaps 0;
 Matches 391; Conservative 0; MisMatches 18; Indels 0; Gaps 0;

QY 391 CTGTTAAACCTGGCTCCGAAATTCAAAGTCACAAACAAACTCTAGCAGCTC 450
 Db 1105 CTGTTAAAAGCTGGCTCCGAAATTCAAAGTCACAAACAAACTCTAGCAGCTC 1164
 QY 451 CTATTAAGAGGAACCCCTGTTCTTATGTAAGTAACTGTTGTAACGTGAAATGGGA 510
 Db 1165 CTATTAAGAGGAACCCCTGTTCTTATGTAAGTAACTGTTGTAACGTGAAATGGGA 1224
 QY 511 GCCAAGAACCCAGAAACCCAGAAACCCGCTCTAGTGACACCCACCAAGACTCTCCACTC 570
 Db 1225 GCCAAGAACCCAGAAACCCAGAAACCCGCTCTAGTGACACCCACCAAGACTCTCCACTC 1284
 QY 571 CCCAAAGGGCCACCCACCCATCCAAAGGGAAGACCTCTCAAGTGAGATAAGATT 630
 Db 1285 CCCAAAGGGCCACCCACCCATCCAAAGGGAAGACCTCTCAAGTGAGATAAGATT 1344
 QY 631 TGCCGTAACCTTCCACCAACAGAAATGTTATGTTGTTGTTGACACCACGCTCCCATC 690
 Db 1345 TGCCGTAACCTTCCACCAACAGAAATGTTATGTTGTTGTTGACACCACGCTCCCATC 1404
 QY 691 CGTACCATTAAGGAAATCTCTCCAAAGAGGAGGACTGAGCAG 739
 Db 1405 CGTTACCATTAAGGAAATCTCTCCAAAGAGGAGGACTGAGCAG 1453
 Search completed: July 21, 2003, 13:29:34
 Job time : 3 secs



1 **NAME_SEQUENCE** 1.0
 LOCUS RSM80235 1433 bp mRNA linear PRI 18-FEB-2000
 DEFINITION Homo sapiens mRNA; cDNA DKFZp586B1922 (from clone DKFZp586B1922).
 ACCESSION AL049450
 VERSION AL049450.1 GI:4500236
 KEYWORD SOURCE
 ORGANISM Homo sapiens (human)
 EUKARYOTA
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1433)
 AUTHORS Koehler, K., Bever, A., Meves, H. W., Gassenhuber, J. and Wiemann, S.
 TITLE Direct Submission
 JOURNAL Submitted (10-MAR-1999) MIFPS, Am Klopferspitz 18a, D-82152
 COMMENT
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.
 This clone (DKFZp586B1922) is available at the RZPD, Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Chaillotzenburg, GERMANY. Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://www.mips.biochem.mpg.de/proj/cDNA/>.
 FEATURES
 source
 1. 1433
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="RZPD:DKFZp586B1922"
 /clone="DKFZp586B1922"
 /tissue_type="uterus"
 /clone_id="586 (synonym: hute1)". Vector psport1; host DH10B; sites NotI + Sall/MuI"
 /dev_stage="adult"
 1386..1391
 POLYA_SIGNAL
 1412
 BASE_COUNT 414 a 321 c 331 g 367 t
 ORIGIN

AL049450 Length: 1433 July 21, 2003 13:18 Type: N Check: 450 ..
 1 TAGGAAGAG CCTCTAACAC AGAACACCT AAGTATCCCT ACTCCCTTC
 51 AAGACTAA TCCCTCAGA GATTCCTT GCTTCAGTC TTCAAGGCC
 101 TCCGAAGAT GCTAGGAGTC GTTAAAGAG GCTGAGAT GGCGGTGCG
 151 AAGGAACTGT TGTAGCTGG CAAACTTCA TCCCAGAG GGGGGCGAA
 201 TACCTCCCTG GCAGCAGAC ATCATGGATT TTCCGCCCTC TGGAAACAG
 251 AGGTCCTCC GTTCTGGTT TGAATCCCT AGAGACAAG AACATCATTC
 301 TCTCCCATAC CGCGGTGGGA CTGCTGGGA GGAGAGGGG TGGAGGGTT
 351 AGACTCTAC CCCATTTGG AAGTACAGAA AGAAAGACTC CTGTAAMA
 401 GCTGGCTCT GATTTCAA AGTCAAAC AAAACTCTC AAGACTCTC
 451 CTATTAAGA GGRACCCGT GTTCTCTT CTGAAACTGT TTGTAACGT
 501 GAATTGAGGA GCGAAGAAC CCCAGAAAG CCCCGCTCT CAGTGACAC
 551 CCCACCAAGA CTCCTCCACTC CCCAAAGGG ACCCAGCACC CATCCAAAG
 601 AGAAGCCTT CTCAGTGAG ATAGAAGATT TGCGTACCT TTCCACACAA
 651 GAATGTATT TGTGTCGTT GCACAGCCT CCCCATCAC CGTACACATT
 701 ACCGAAATCC TCTCCAAAGA AGGAGGAGAC TGTAGCAAGT AAGGCATAGA

751 GAAACTTCG TCTTATACCC TAGTGGGCC GGTCAAGCTA ACAGGTGTA
 801 AAATGCCCTT GCGATTTTA AAAAGTCA ATCATTAAG CAGAGCTCTG
 851 TCAAGAATGA GTAAAGTTAC AGCCAGAGAC AGACACTGTG CAGGATTGC
 901 AAATAGATGG AATTACAGCA AATGTGCTC AATGTATTG CCTGGTTACA
 951 ACATGGGGAG ATGTGTTGC CAGTAAGTG CTCATACAA GAGCAGCAGA
 1001 CTTGGGGGTG TAATCTCCCG CAACTTCAT GCCCTTGAA AGAAGGGTT
 1051 TCTGTGCTGT GAATGCTATA GAACTATCT TTGCCATGCA CGACTGTCC
 1101 TGCATTGAT ATTGTGCAA ATCTGGAGG GTGGTCTTG GGTGTTCTCA
 1151 GGGCCAAATG GTAAATTG GCTTGGGGAG CCAGCTGGG GTGGGGAATT
 1201 TTCACTGGG CCTCCGCCTT TTRACTAT AACATTTT CTGTATATCT
 1251 ATGTCCTGT CTGGGGCA GGAGAATCT GCCAANGACC AACAGCTTA
 1301 CTTTATCTTA CTATACATCA CAAAGTCTT AAAATGTGAA GAGTTTACTT
 1351 GGATGCACT AGCCATTTG TTGTCTAT ATTAAATAA ATGGCTTAC
 1401 AACTATTCT TCAACACAAA AAAAAAAA AAA

